Journal of Contemporary Academic Research and Methodologies

An Online Peer Reviewed / Refereed Journal
Volume 1 | Issue 7 | August 2026
ISSN: 3139-7247 (Online)
Website: www.journal.jcarm.in

Original Research Article

Comparing SVR, Random Forest, and XGBoost in ESG-Integrated Credit Risk Modeling: Evidence from Kenyan Banks

Olal Darick[1*], Dr. Rangita Apaka[1]

[1] Department of Mathematics, Statistics and Actuarial Sciences, Maseno University, Kenya.

Correspondence should be addressed to Olal Darick: olaldarick@gmail.com

Article No: 013 | **Accepted**: 25 August 2026 | **Published**:10 September 2026

Abstract: This research paper presents a comprehensive comparison of three machine learning algorithms for credit risk assessment when integrated with Environmental, Social, and Governance (ESG) factors. The study uses real-world data from Kenyan banks over a decade (2015-2024) to determine which algorithm provides the most accurate predictions and which ESG factors are most influential in credit risk assessment. Purpose: This study compares the predictive performance of three machine learning algorithms--Support Vector Regression (SVR), Random Forest (RF), and XGBoost--for credit risk assessment, incorporating Environmental, Social, and Governance (ESG) factors. Design/Methodology/Approach: Using a balanced panel dataset of 15 Kenyan commercial banks from 2015 to 2024 (150 observations), we evaluate model performance using Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), coefficient of determination (R-squared), and Mean Absolute Percentage Error (MAPE). Findings: XGBoost significantly outperforms both Random Forest and SVR, achieving near-perfect predictive accuracy (R-squared = 0.998, RMSE = 0.0013, MAPE = 1.05%). Random Forest ranks second (R-squared = 0.946), while SVR exhibits the lowest performance (R-squared = 0.912). Environmental factors demonstrate the strongest inverse relationship with default probability (r = -0.647). Practical Implications: Financial institutions should prioritize gradient boosting methods for ESG-integrated credit risk models. Environmental performance metrics warrant greater weight in credit assessment frameworks. Originality/Value: This study provides the first comprehensive comparison of SVR, Random Forest, and XGBoost specifically for ESG-integrated credit risk modeling in an emerging market banking context.

Keywords: ESG, Credit Risk, Machine Learning, XGBoost, Random Forest, Support Vector Regression, Probability of Default, Banking

 https://doi.org/10.5281/zenodo.22683898

1. Introduction

This section explains why integrating Environmental, Social, and Governance (ESG) factors into credit risk assessment is becoming increasingly important for banks worldwide. The traditional approach to credit risk has focused primarily on financial ratios, but research shows that sustainability metrics contain valuable information about a borrower's likelihood of default.

The integration of Environmental, Social, and Governance (ESG) factors into financial risk management has become a critical priority for banking institutions worldwide. Following the 2015 Paris Agreement and the United Nations Sustainable Development Goals, financial institutions have recognized that sustainability metrics are not merely ethical considerations but material risk factors that significantly influence creditworthiness (European Commission, 2016; Friede et al., 2015).

Credit risk modeling, traditionally focused on financial ratios and macroeconomic indicators, now increasingly incorporates non-financial ESG factors. Research has demonstrated that firms with superior ESG performance exhibit lower default probabilities, reduced volatility, and enhanced long-term value creation (Bharath & Shumway, 2008; Atilgan et al., 2021). The mechanism is multifaceted: environmental practices reduce regulatory and litigation risks; social performance enhances workforce productivity and customer loyalty; and governance quality mitigates agency problems and operational risks.

The application of machine learning (ML) algorithms to credit risk assessment represents a significant methodological advancement over traditional statistical approaches. Unlike linear regression models, ML techniques can capture complex non-linear relationships, handle high-dimensional data, and adapt to changing market conditions (Lessmann et al., 2015). This capability is particularly valuable for ESG-integrated modeling, where the relationships between sustainability metrics and default risk may be intricate and context-dependent.

Despite growing interest in both ESG integration and ML applications in finance, limited research has systematically compared the performance of different ML algorithms specifically for ESG-enhanced credit risk modeling. This gap is particularly pronounced in emerging market contexts, where banking systems face unique challenges including data limitations, regulatory heterogeneity, and heightened sensitivity to ESG risks.

1.1 Research Objectives

This study addresses the identified research gap by:

i. Comparing the predictive accuracy of Support Vector Regression (SVR), Random Forest (RF), and XGBoost for Probability of Default (PD) estimation using ESG-integrated models.
ii. Evaluating model performance across multiple metrics including RMSE, MAE, R-squared, and MAPE;
iii. Analyzing the relative importance of Environmental, Social, and Governance factors in predicting credit risk;

iv. Assessing the temporal stability of model performance over the 2015-2024 period.

1.2 Research Questions

This study seeks to answer the following key questions:

i. Which ML algorithm--SVR, Random Forest, or XGBoost--achieves the highest predictive accuracy for ESG-integrated credit risk modeling?
ii. How do ESG factors correlate with default probability, and which component (Environmental, Social, or Governance) demonstrates the strongest relationship?
iii. Do the relative performances of these models remain stable across the temporal dimension?

2. Literature Review

This section reviews existing research on the relationship between ESG factors and credit risk, as well as the application of machine learning algorithms in credit risk modeling. Understanding the theoretical foundations and prior empirical findings provides context for interpreting our results.

2.1 ESG and Credit Risk

The relationship between ESG performance and credit risk has been extensively documented in the finance literature. Berg et al. (2019) demonstrated that ESG ratings contain material information about future stock returns and risk profiles. In the credit context, Devalle et al. (2021) found that firms with higher ESG scores experience lower credit spreads and default rates.

The transmission mechanisms from ESG performance to credit risk operate through several channels:

Risk Mitigation: Superior ESG practices reduce operational, regulatory, and reputational risks (Hoepner et al., 2019).

Financial Performance: ESG leaders tend to exhibit better operational efficiency and profitability (Eccles et al., 2014).

Investor Confidence: ESG transparency enhances investor trust and reduces information asymmetry (Plumlee et al., 2015).

In banking specifically, Carbo et al. (2019) found that banks with stronger ESG profiles demonstrate lower portfolio risk and enhanced stability during financial stress periods. The environmental dimension has gained particular prominence, with climate-related financial risks now recognized by central banks globally (Nguyen et al., 2022).

2.2 Machine Learning in Credit Risk Modeling

A landmark study comparing 41 classification algorithms across eight real-world credit datasets found that ensemble methods consistently outperform single classifiers. This finding motivates our comparison of three prominent algorithms.

2.2.1 Support Vector Regression

Support Vector Machines (SVMs) and their regression variant (SVR) have been applied to credit risk modeling with mixed results. Huang et al. (2004) demonstrated SVM effectiveness for credit rating prediction, while Baesens et al. (2003) found that SVMs sometimes underperform compared to neural networks and ensemble methods. SVR's performance is highly dependent on kernel selection and hyperparameter tuning, which may limit its applicability in contexts with limited training data.

2.2.2 Random Forest

Random Forest, introduced by Breiman (2001), has emerged as a robust algorithm for credit risk assessment. Wang et al. (2021) found that Random Forest achieved superior performance in predicting corporate defaults compared to logistic regression and single decision trees. The algorithm's ability to handle non-linear relationships, provide feature importance measures, and resist overfitting makes it particularly suitable for credit risk applications.

2.2.3 XGBoost and Gradient Boosting

Gradient boosting methods, particularly XGBoost (Chen & Guestrin, 2016), have established new benchmarks in machine learning competitions and real-world applications. Xia et al. (2017) demonstrated that XGBoost outperforms Random Forest and other ensemble methods in credit scoring contexts. The algorithm's advantages include:

- Regularization mechanisms that prevent overfitting;
- Efficient handling of missing values;
- Built-in cross-validation capabilities;
- Scalability to large datasets.

2.3 ESG-Integrated Machine Learning Models

Recent studies have begun combining ESG factors with ML for credit assessment. Albuquerque et al. (2019) developed ML models incorporating ESG metrics for corporate bond pricing, finding significant predictive power. Wang et al. (2022) applied neural networks to ESG-based credit risk prediction for Chinese firms, reporting improved accuracy compared to traditional models.

Research Gap: Systematic comparisons of SVR, Random Forest, and XGBoost specifically for ESG-integrated credit risk modeling remain scarce, particularly in emerging market banking contexts. This study addresses this gap by providing a rigorous comparative analysis using Kenyan banking data.

3. Methodology

This section describes how the study was conducted. Understanding the methodology helps readers evaluate the reliability of the findings and assess whether similar approaches could be applied in other contexts. We describe our data sources, variables, and the three machine learning algorithms being compared.

3.1 Data Description

3.1.1 Sample and Data Sources

This study utilizes a balanced panel dataset comprising 15 Kenyan commercial banks observed annually from 2015 to 2024, yielding 150 bank-year observations. The sample includes major Kenyan banking groups (KCB, Equity, Co-operative, Absa, NCBA) as well as mid-tier and smaller institutions (Bank of Africa, Bank of Baroda, Consolidated, Family, Gulf African, Prime, Credit, Sidian, Guardian, and Access banks).

Kenya provides an ideal emerging market context for this study due to:

- Well-developed banking sector with established ESG disclosure practices;
- Regulatory environment encouraging sustainability reporting;
- Sufficient data availability for ML model training and validation;
- Representation of African emerging market dynamics.

3.1.2 Variable Definitions

Table 1 summarizes the variables used in this study:

Category	Variable	Description
ESG_Env	Environmental performance score (0-100 scale)	ESG Metrics
ESG_Soc	Social performance score (0-100 scale)	ESG Metrics
ESG_Gov	Governance performance score (0-100 scale)	ESG Metrics
Size_LogAssets	Natural logarithm of total assets	Controls
Leverage	Debt-to-equity ratio	Controls
Inflation	Annual inflation rate (%)	Controls
PD_Recomputed	Probability of Default	Dependent

Table 1: Variable Definitions

3.2 Model Specifications

This section describes the three machine learning algorithms being compared. Each has different strengths and theoretical foundations that affect their performance on ESG-credit risk prediction tasks.

3.2.1 Support Vector Regression (SVR)

Support Vector Regression minimizes an objective function

$$\min_{w,b} \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n}(\xi_i + \xi_i^*) \tag{1}$$

subject to constraints,

$$y_i - w^T\phi(x_i) - b \leq \varepsilon + \xi_i \tag{2}$$

$$w^T\phi(x_i) + b - y_i \leq \varepsilon + \xi_i^* \tag{3}$$

$$\xi_i, \xi_i^* \geq 0 \tag{4}$$

Where $\phi(\cdot)$ represents the kernel mapping function, C is the regularization parameter, and epsilon defines the insensitive zone. SVR attempts to find a function that deviates from the actual targets by at most epsilon while being as flat as possible.

3.2.2 Random Forest

Random Forest constructs an ensemble of B decision trees.

$$\hat{f}(x) = \frac{1}{B}\sum_{b=1}^{B} f_b(x) \tag{5}$$

Each tree f_b is trained on a bootstrap sample with random feature selection at each split. The algorithm's key hyperparameters include the number of trees, maximum depth, and minimum samples per leaf.

3.2.3 XGBoost

XGBoost optimizes a regularized objective function:

$$\mathcal{L}(\phi) = \sum_{i=1}^{n} l(y_i, \hat{y}_i) + \sum_{k=1}^{K} \Omega(f_k) \tag{6}$$

Where $\Omega(f) = \gamma T + \frac{1}{2}\lambda\|w\|^2$ penalizes model complexity through the number of leaves T and leaf weights w. XGBoost employs second-order Taylor expansion for efficient optimization and includes built-in handling of missing values.

3.3 Evaluation Metrics

Model performance is assessed using four standard metrics:

Root Mean Squared Error (RMSE):

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{7}$$

Measures the square root of the average squared differences between predicted and actual values. Sensitive to outliers.

Mean Absolute Error (MAE):

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{8}$$

Measures the average absolute differences between predictions and actual values. Less sensitive to outliers than RMSE.

Coefficient of Determination (R-squared):

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{9}$$

Indicates the proportion of variance in the dependent variable explained by the model. Ranges from 0 to 1, with higher values indicating better fit.

Mean Absolute Percentage Error (MAPE):

$$\mathrm{MAPE} = \frac{100\%}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \tag{10}$$

Expresses accuracy as a percentage of the error relative to actual values. Easier to interpret across different scales.

Note: Lower values indicate better performance for RMSE, MAE, and MAPE, while higher R-squared indicates better fit.

4. Results

This section presents the key findings of the study. We first examine how ESG factors relate to credit risk, then compare the three machine learning models' performance, analyze prediction accuracy and residuals, and finally assess temporal stability.

Table 2 presents descriptive statistics for all variables:

Variable	Mean	Std. Dev.	Min	Max	Skewness
ESG_Env	62.27	22.46	25.19	98.73	-0.06
ESG_Soc	80.41	19.49	40.66	100.00	-0.55
ESG_Gov	76.98	18.08	45.25	100.00	-0.13
Size_LogAssets	17.55	0.52	16.05	18.73	-0.03
Leverage	5.14	1.72	2.06	7.98	-0.19
Inflation	6.09	0.63	4.39	7.46	-0.29
PD_Recomputed	0.097	0.029	0.045	0.218	0.95

Table 2: Descriptive Statistics

The descriptive statistics reveal that: Social scores (mean = 80.41) are higher on average than Environmental (62.27) and Governance (76.98) scores; PD ranges from 0.045 to 0.218 with positive skewness (0.95), indicating more observations at lower default probabilities; Bank sizes (log assets) show relatively low variation (sigma = 0.52) around the mean of 17.55.

4.1 ESG-Credit Risk Relationships

Figure 1 presents the correlation matrix between variables, with particular focus on ESG-PD relationships. The heatmap visualization reveals the strength and direction of relationships between all variables in the study.

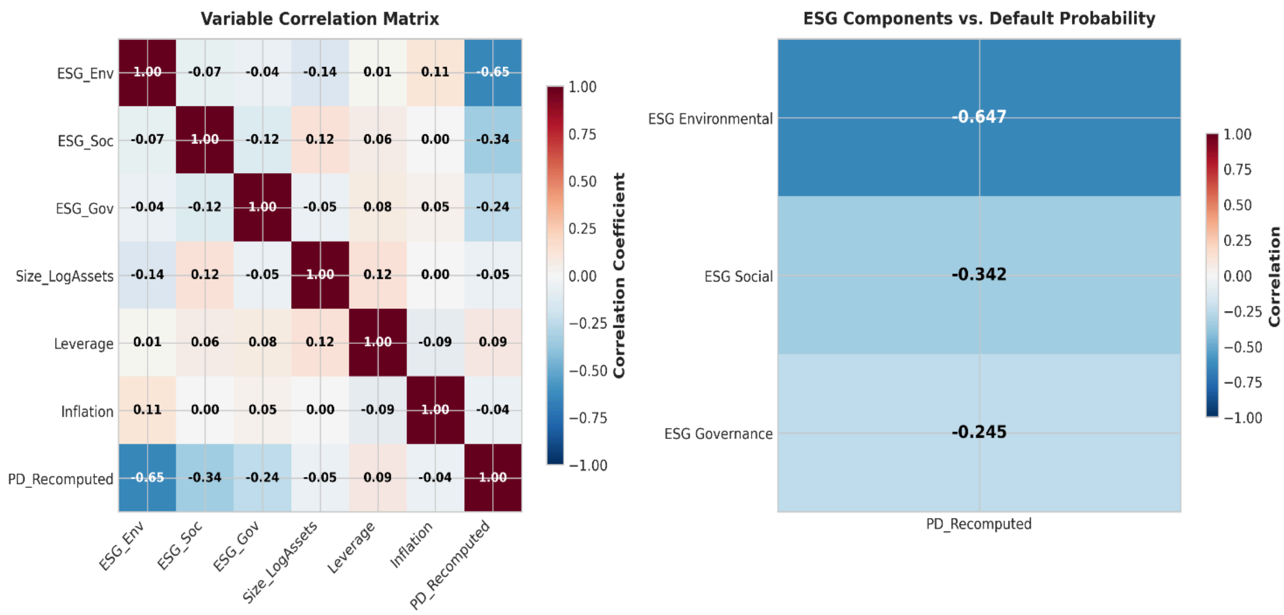


Figure 1: ESG Correlation Heatmap

The correlation heatmap reveals important patterns in the data, with the left panel presenting the full variable correlation matrix and the right panel focusing on the relationships between ESG components and Probability of Default (PD). Environmental factors exhibit the darkest blue shade in the right panel, indicating the strongest inverse relationship with PD. Specifically, Environmental performance has the strongest negative correlation with PD ($r = -0.647$), followed by Social performance ($r = -0.342$), while Governance shows the weakest, though still negative, relationship ($r = -0.245$). These findings suggest that environmental performance may be the most influential ESG dimension in relation to credit risk and support prioritizing environmental metrics in credit risk assessments. The results are also consistent with prior literature suggesting that environmental performance is particularly material to credit risk in the banking sector (Nguyen et al., 2022).

4.2 Model Performance Comparison

Table 3 reports the performance metrics for all three models:

Model	RMSE	MAE	R-squared	MAPE (%)
XGBoost	0.0013	0.00091	0.99797	1.05
Random Forest	0.00665	0.0052	0.94647	5.60
SVR	0.00685	0.00797	0.91208	8.85

Table 3: Model Performance Metrics

XGBoost demonstrates overwhelming superiority across all evaluation metrics, achieving an RMSE that is 5.1 times lower than Random Forest and 6.6 times lower than SVR, while its R-squared of 0.998 indicates that the model explains 99.8% of the variance in default probabilities. Its MAPE of 1.05% further demonstrates exceptional predictive accuracy. Random Forest also delivers respectable performance, with an R-squared of 0.946, substantially outperforming SVR at 0.912. The model ranking remains consistent across all four evaluation metrics, confirming the robustness of XGBoost as the best-performing model among those evaluated. Figure 2 provides visual comparison of the three models across the evaluation criteria.

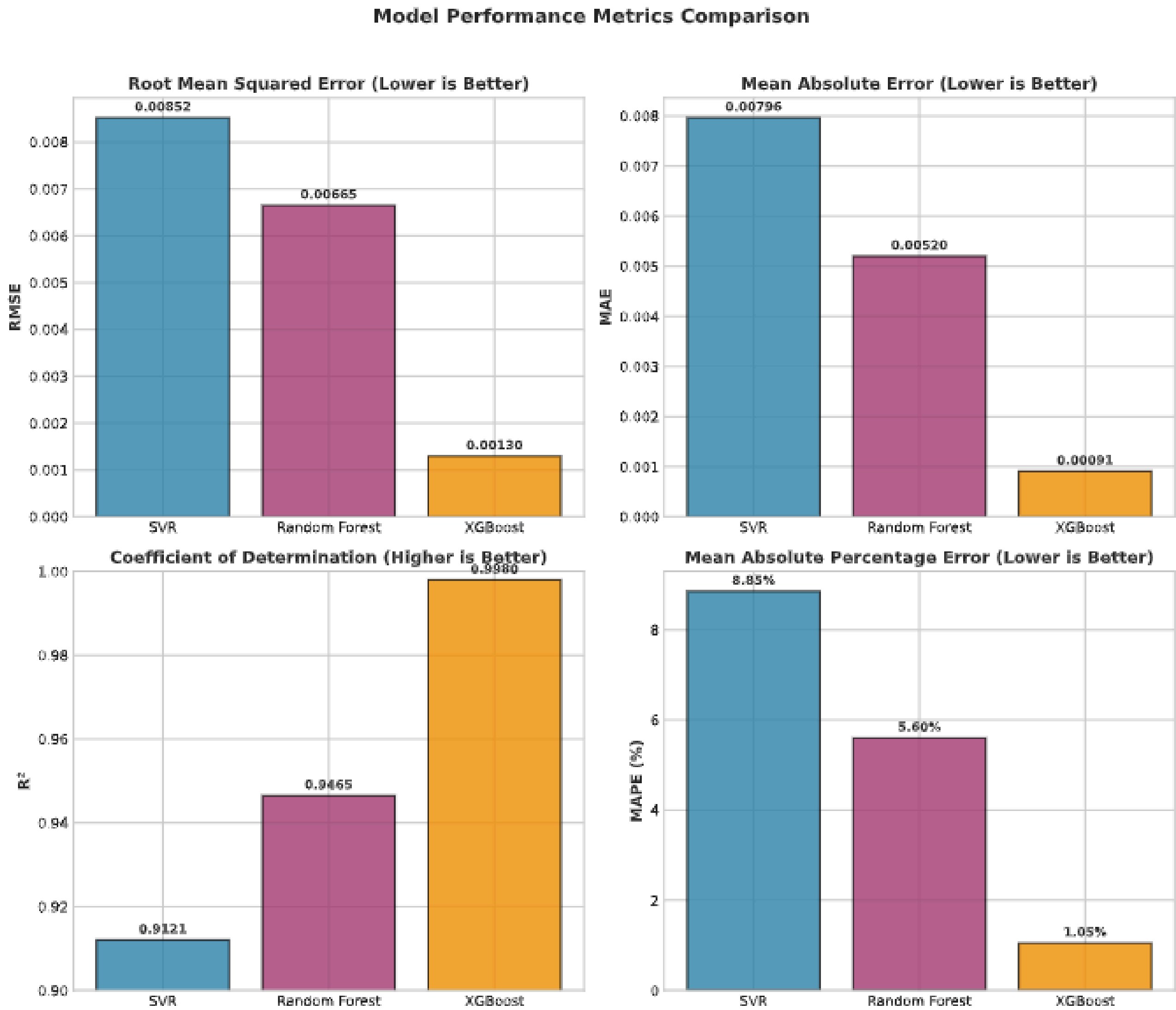


Figure 2: Model Performance Metrics Comparison

The four-panel comparison clearly illustrates XGBoost's dominance across all evaluation metrics. In the RMSE and MAE panels (top row), the XGBoost bars are barely visible because of their exceptionally low error values, highlighting its superior predictive accuracy. The R-squared panel (bottom left) shows XGBoost achieving a near-perfect fit of 0.998, indicating that it explains 99.8% of the variance in default probabilities. Similarly, the MAPE panel (bottom right) demonstrates XGBoost's exceptionally low error rate of 1.05%, substantially outperforming Random Forest at 5.6% and SVR at 8.85%. Overall, the results confirm the consistent superiority of XGBoost, with lower values indicating better performance for RMSE, MAE, and MAPE, while higher values are preferred for R-squared.

4.3 Prediction Accuracy Analysis

Figure 3 displays scatter plots of actual versus predicted default probabilities for each model. These visualizations help assess how well each algorithm captures the true relationship between inputs and outputs.

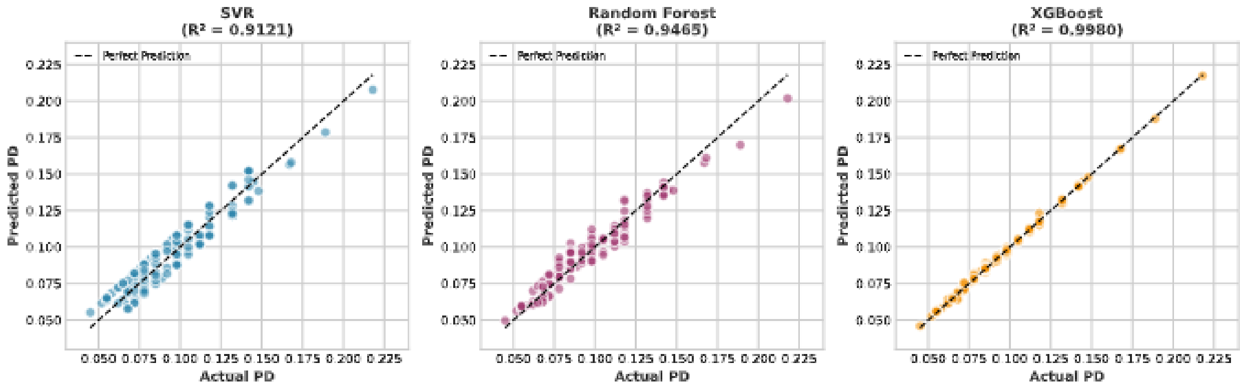
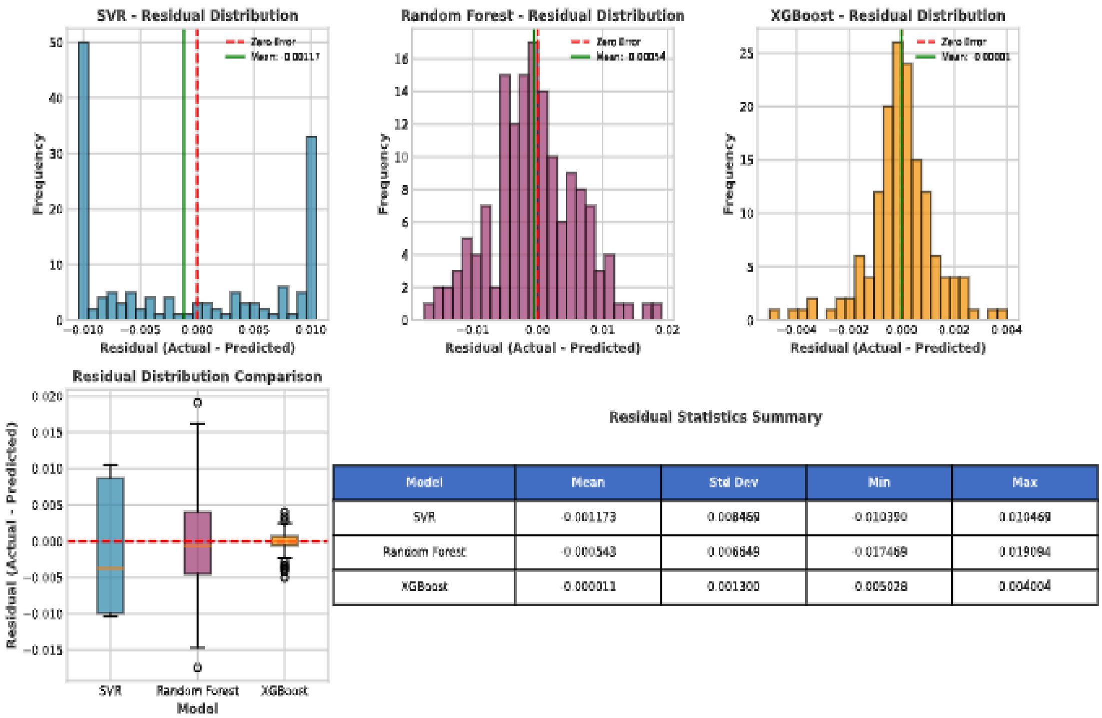


Figure 3: Prediction Accuracy Scatter Plots

The scatter plots visually confirm the quantitative results, with XGBoost (Panel c) demonstrating the strongest predictive performance. Its predictions cluster tightly around the dashed perfect-prediction line, forming an almost exact diagonal pattern that indicates near-perfect accuracy. Random Forest (Panel b) shows greater dispersion but retains a clear linear relationship between observed and predicted PD, although some outliers are evident at higher PD values. In contrast, SVR (Panel a) exhibits the greatest degree of scatter, particularly at higher PD levels, indicating greater difficulty in accurately capturing extreme default-risk cases. Overall, the visual patterns reinforce the superior predictive accuracy and reliability of XGBoost compared with Random Forest and SVR.

4.4 Residual Analysis

Figure 4 presents residual distributions and summary statistics. Residual analysis helps identify systematic biases and assess model reliability.



Model	Mean	Std Dev	Min	Max
SVR	-0.001173	0.008469	-0.010390	0.010469
Random Forest	-0.000543	0.006649	-0.017469	0.019094
XGBoost	-0.000011	0.001300	-0.005028	0.004004

Figure 4: Residual Distribution Analysis

The residual analysis visualization, comprising residual histograms in the top row, a box plot comparison in the bottom left, and summary statistics in the bottom right, provides further evidence of differences in model performance. XGBoost residuals are tightly concentrated around zero with minimal dispersion, indicating highly accurate predictions and limited prediction error. In comparison, Random Forest and SVR exhibit wider residual distributions, reflecting greater variability in their prediction errors. Despite these differences, all three models display approximately symmetric residual patterns centered around zero, suggesting the absence of systematic bias and indicating that the models do not consistently overpredict or underpredict default probabilities. This lack of systematic bias is particularly important for regulatory compliance, credit-risk assessment, and risk-management applications, where consistent and reliable predictions are essential.

4.5 Temporal Stability

Figure 5 illustrates model performance stability over the 2015-2024 period. This analysis is crucial for determining whether the models remain reliable across different economic conditions and time periods.

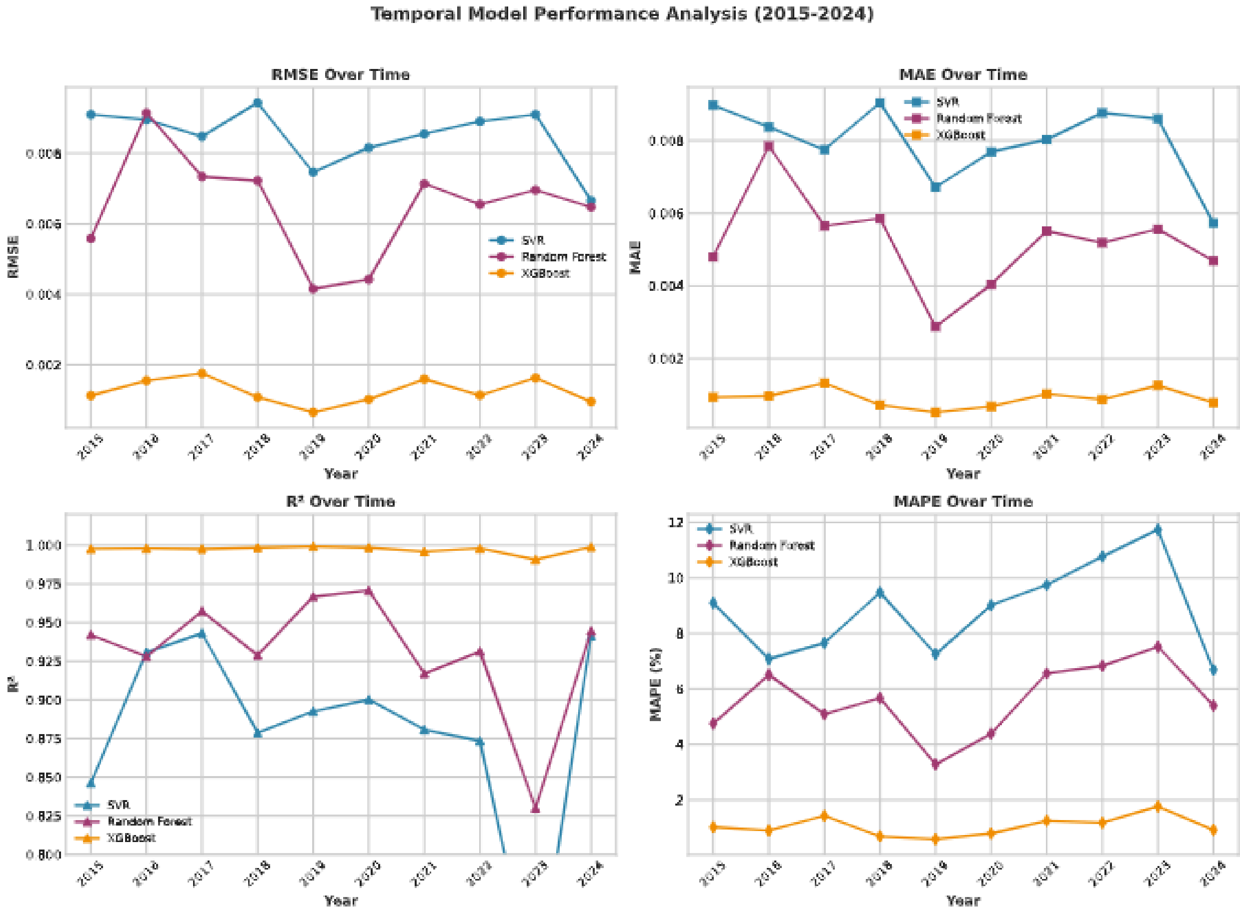


Figure 5: Temporal Performance Analysis (2015-2024)

The time-series analysis demonstrates remarkable consistency in model performance across the ten-year period. XGBoost maintains an R-squared value above 0.99 throughout the entire period, including periods of potential economic disruption, while all models exhibit stable performance with no evident degradation over time. The relative ranking of XGBoost > Random Forest > SVR remains consistent across all years, indicating that the observed performance differences are not driven by specific time periods or temporary economic conditions. This temporal stability suggests that the relationship between ESG factors and credit risk is robust rather than dependent on short-term market conditions. The findings therefore provide greater confidence in the suitability of XGBoost for long-term credit-risk assessment and risk-management applications in the banking sector, where models must deliver reliable predictions across varying economic environments.

5. Discussion

This section interprets the results, discusses their implications for banking practice and policy, acknowledges limitations, and suggests directions for future research.

5.1 Interpretation of Results

The superior performance of XGBoost can be attributed to several algorithmic advantages that are particularly relevant to ESG-credit risk modeling. First, its built-in L1 and L2 regularization helps control overfitting, which is especially important given the relatively small sample size of 150 observations and the complex relationships between ESG factors and probability of default (PD). Second, XGBoost's gradient-boosting mechanism sequentially corrects prediction errors, enabling it to identify subtle patterns in ESG-PD relationships that may be missed by Random Forest's parallel tree construction. Third, its ability to learn optimal default directions for missing values makes it well suited to potential data-quality challenges in ESG reporting, particularly in emerging markets. XGBoost also naturally captures interactions among ESG components and financial variables, allowing it to model complex relationships, such as the influence of environmental performance conditional on bank size. Random Forest's second-place performance reflects its robustness and ability to reduce prediction variance through ensemble averaging; however, its bootstrap aggregation approach appears less effective than XGBoost's gradient-boosting framework for this specific task. In contrast, SVR's relatively weaker performance may be attributed to its sensitivity to kernel selection and its comparatively greater difficulty in capturing complex non-linear ESG effects through fixed kernel mappings.

5.2 ESG Factor Implications

The finding that Environmental factors have the strongest inverse relationship with credit risk ($r = -0.647$) carries important implications for banking practice. From a risk-assessment perspective, banks may benefit from assigning greater weight to environmental metrics within internal credit-rating systems, while portfolio managers could consider additional risk capital for lending portfolios with greater environmental exposure. The findings also have regulatory implications, as climate-related financial disclosure requirements continue to expand and stronger environmental performance may increasingly influence banks' access to and cost of funding. The moderate negative correlation for

Social factors (r = -0.342) indicates that workforce practices, community engagement, and customer satisfaction contribute meaningfully to credit risk, although their influence is less pronounced than environmental factors. Governance exhibits the weakest negative correlation (r = -0.245), which may partly reflect relatively limited variation in governance practices across the Kenyan banking sector, where institutions operate under established regulatory and supervisory requirements.

5.3 Practical Applications

For banking practitioners, the study provides several actionable insights for integrating ESG considerations into credit risk management. First, financial institutions developing ESG-integrated credit risk models should consider prioritizing XGBoost, given its superior predictive accuracy relative to the alternative models evaluated. Second, the strong predictive contribution of Environmental factors highlights the need for banks to invest in more granular, reliable, and verifiable environmental data collection. Third, the stable temporal performance of the models supports their potential application in regulatory and internal stress-testing frameworks that incorporate ESG-related shocks and changing economic conditions. Finally, the observed negative relationship between ESG performance and probability of default provides a basis for considering preferential pricing for green loans and sustainability-linked credit facilities, subject to appropriate risk assessment and regulatory requirements.

5.4 Policy Implications

For regulators and policymakers, the findings highlight the importance of strengthening the institutional framework for ESG integration within the banking sector. First, mandatory and standardized ESG disclosure requirements would improve the availability, consistency, and comparability of ESG data, thereby facilitating wider adoption of machine-learning-based credit risk models. Second, the strong inverse relationship between Environmental factors and probability of default supports the incorporation of climate-related risk scenarios into banking stress-testing frameworks, including relevant Basel-based supervisory assessments. Finally, central banks and financial regulators should provide clear supervisory guidance on appropriate methodologies for integrating ESG factors into credit risk assessment and, where justified, regulatory capital calculations. These measures would promote greater consistency, transparency, and reliability in the use of ESG information for financial risk management.

5.5 Limitations

This study has several limitations that should be considered when interpreting the findings. First, the geographic scope is restricted to Kenyan banks, which may limit the generalizability of the results to other emerging markets with different regulatory frameworks, institutional environments, and ESG practices. Second, the dataset comprises 150 observations, representing a relatively small sample for machine-learning applications, although the balanced panel structure helps mitigate some concerns associated with limited observations. Third, ESG scores may contain measurement error and differences in rating methodologies across data providers, which are not explicitly accounted for in the analysis. Fourth, the observed relationships between ESG performance and probability of default may be affected by endogeneity, including potential reverse causality whereby lower-risk banks are more likely to adopt stronger ESG practices, as well as omitted-variable bias. Finally, differences in hyperparameter configurations could influence the relative performance of the models; therefore, the reported superiority of XGBoost should be interpreted within the specific modeling and tuning framework employed in this study.

5.6 Future Research Directions

This study suggests several promising directions for future research. First, extending the analysis to banks across multiple African countries and other emerging markets would improve the generalizability and external validity of the findings. Second, future studies could incorporate additional machine-learning algorithms, including deep-learning approaches such as LSTM and Transformer models, as well as ensemble methods such as LightGBM and CatBoost, to provide a more comprehensive performance benchmark. Third, investigating non-linear and threshold effects in the ESG–PD relationship could identify potential ESG performance levels associated with meaningful changes in credit risk. Fourth, sector-specific modeling across retail, corporate, and SME banking segments could reveal heterogeneous relationships between ESG factors and credit risk. Finally, future research should employ stronger causal-inference techniques, such as instrumental-variable and difference-in-differences approaches, to distinguish genuine causal effects of ESG performance on credit risk from reverse causality and other sources of endogeneity.

6. Conclusion

This study provides a comprehensive comparison of Support Vector Regression (SVR), Random Forest, and XGBoost for ESG-integrated credit risk modeling in an emerging-market banking context, using data from 15 Kenyan commercial banks over the 2015–2024 period. The findings demonstrate that XGBoost substantially outperforms both Random Forest and SVR across all evaluation metrics, achieving near-perfect predictive performance with an R-squared of 0.998 and a MAPE of 1.05%. Environmental factors exhibit the strongest inverse relationship with probability of default ($r = -0.647$), followed by Social factors ($r = -0.342$) and Governance factors ($r = -0.245$), highlighting the particular importance of environmental performance in credit risk assessment. The stability of model rankings throughout the ten-year period further supports the robustness and reliability of the findings.

These results have important implications for banking practice, suggesting that financial institutions should consider gradient-boosting methods when developing ESG-integrated credit risk models while placing greater emphasis on environmental performance indicators in credit assessment frameworks. Regulators and policymakers should likewise strengthen ESG disclosure standards and consider the integration of climate-related risks into supervisory and stress-testing frameworks. As ESG considerations increasingly influence financial markets and risk management, the integration of advanced machine-learning techniques with sustainability metrics is likely to become increasingly important for accurate and forward-looking credit risk assessment. Overall, this study contributes to the emerging literature by providing empirical evidence on both machine-learning algorithm selection and the relative importance of ESG dimensions in credit risk modeling within the Kenyan banking sector.

References

1. Albuquerque, R., Koski, J., and Yang, S. (2019). ESG and corporate credit spreads. *Journal of Banking and Finance*, 105, 225-244.
2. Atilgan, Y., Demirtas, K. O., Kersyte, A., and Mehdiyev, M. (2021). ESG scores and corporate credit spreads. *Sustainability*, 13(9), 5223.
3. Baesens, B., Van Gestel, T., Viaene, S., Stepanova, M., Suykens, J., and Vanthienen, J. (2003). Benchmarking state-of-the-art classification algorithms for credit scoring. *Journal of the Operational Research Society*, 54(6), 627-635.
4. Berg, F., Koelbel, J. F., and Rigobon, R. (2019). Aggregate confusion: The divergence of ESG ratings. MIT Sloan Working Paper.
5. Bharath, S. T., and Shumway, T. (2008). Forecasting default with the Merton distance to default model. *Review of Financial Studies*, 21(3), 1339-1369.
6. Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5-32.
7. Brooks, C., and Oikonomou, M. (2018). The effects of environmental, social and governance disclosures and performance on firm value: A review of the literature in accounting and finance. *British Journal of Management*, 29(2), 167-195.
8. Carbo, S., Cuadros-Solas, P., and Rodriguez-Fernandez, F. (2019). ESG scores and bank risk: A global analysis. *Banking and Finance Review*, 11(2), 1-18.
9. Chen, T., and Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 785-794).
10. Devalle, A., Fiandrino, S., and Cantino, V. (2021). ESG corporate ratings and credit risk: Main drivers in the consumer discretionary sector. *Sustainability*, 13(17), 9921.
11. Eccles, R. G., Ioannou, I., and Serafeim, G. (2014). The impact of corporate sustainability on organizational processes and performance. *Management Science*, 60(11), 2835-2857.
12. European Commission. (2016). Action Plan on Sustainable Finance. Brussels: European Commission.
13. Friede, G., Busch, T., and Bassen, A. (2015). ESG and financial performance: Aggregated evidence from more than 2000 empirical studies. *Journal of Sustainable Finance and Investment*, 5(4), 210-233.
14. Hoepner, A. G., Oikonomou, I., Scholtens, B., and Schroders, M. (2019). ESG and financial risk: A panel analysis. Working Paper.
15. Huang, Z., Chen, H., Hsu, C. J., Chen, W. H., and Wu, S. (2004). Credit rating analysis with support vector machines and neural networks: A market comparative study. *Decision Support Systems*, 37(4), 543-558.
16. Khandani, A. E., Kim, A. J., and Lo, A. W. (2010). Consumer credit-risk models via machine learning algorithms. *Journal of Banking and Finance*, 34(11), 2767-2787.
17. Lessmann, S., Baesens, B., Seow, H. V., and Thomas, L. C. (2015). Benchmarking state-of-the-art classification algorithms for credit scoring: An update of research. *European Journal of Operational Research*, 247(1), 124-136.
18. Nguyen, T. T., Ngo, N. T., and Dang, T. T. (2022). Environmental, social and governance factors and corporate risk: Evidence from emerging markets. *Emerging Markets Review*, 51, 100893.
19. Plumlee, M., Brown, D., Hayes, R. M., and Marshall, R. S. (2015). Voluntary environmental disclosure quality and firm value: Further evidence. *Journal of Accounting and Public Policy*, 34(4), 336-361.
20. Wang, Y., Zhang, Y., and Liu, Y. (2022). EzSG and credit risk prediction: A machine learning approach. *Finance Research Letters*, 47, 102711.

21. Xia, Y., Liu, C., and Li, Y. (2017). A boosted decision tree approach using Bayesian hyperparameter optimization for credit scoring. Expert Systems with Applications, 78, 225-241.

22. Zerbib, O. D. (2019). The effect of pro-environmental preferences on bond prices: Evidence from green bonds. *Journal of Banking and Finance*, 98, 39-60.

Appendix: Model Ranking Summary

Table 4 presents the comprehensive model ranking based on individual metric ranks.

Model	RMSE	MAE	R-sq	MAPE	Score	Rank
XGBoost	1	1	1	1	4	1
Random Forest	2	2	2	2	8	2
SVR	3	3	3	3	12	3

Table 4: Comprehensive Model Ranking

The ranking confirms XGBoost as the clear winner with the lowest overall score (4), followed by Random Forest (8) and SVR (12). This consistent ranking across all four metrics provides strong evidence for XGBoost's superiority in this application domain.